**2nd International Working Group Meeting - CHAD project**

The international working group, formed in the frame of the CHAD project, had its second meeting on 14 December 2022. The meeting aimed **to gain an overview of monitoring hate speech online** by getting to know the monitoring methods, experiences, and results of external guests and some participating organisations. After a general introduction of the project and an introductory round, four presentations were held on monitoring hate speech online. A short discussion on further experiences and thoughts on the topic followed these.

**Participants of the second meeting** included people from across Europe working at organisations dealing with hate speech monitoring. Ten participants were present, three from Hungary, one from Poland, Slovakia, Slovenia, Italy, Belgium and the UK, and one representing a global organisation.

<u>**Summary of the presentations on monitoring hate speech online**</u>

***Hate crime, hate speech, and global community tensions – UK***
Initially, the presenter provided a brief theoretical background on why monitor hate crime and hate speech is essential. They then introduced a group bringing together several civil, governmental and academic organisations from the UK to fight against hate speech. The group's "toolbox" consists of enforcement, "traditional" intelligence methods, analysis, engagement with industry and stakeholders, public communication campaigns, and counter-narrative operations. They also explained how their tool for monitoring hate speech works in practice.

***Antisemitic hate speech online – experiences of a global organisation***
The organisation represents the interests of several Jewish communities on an international and diplomatic level, it gets to know about hate speech cases from its partner organisations. The presenter introduced the various forms of online antisemitic hate speech (comments, blog posts, images, etc.) and the possible underlying ideologies (neo-Nazi propaganda, conspiracy myths, etc.). The organisation's further activities include identifying emerging phenomena, policy level impact on social media platforms, de-platforming of individuals, thus disrupting their activities, and disrupting the sale of Holocaust-Denial items. They also spoke about the current Austrian and German situation regarding regulating specific social media platforms and the benefits of de-platforming. Lastly, they outlined some challenges concerning hate speech, such as finding the balance between freedom of expression and the removal of hateful content, the phenomenon that the content goes viral quicker than it could be removed, etc. Furthermore, they presented some successes concerning removing online Holocaust denial content.

***The experience of a Hungarian organisation with monitoring anti-LGBTQI hate speech online***
The participant briefly presented their organisation in general, particularly its experience monitoring anti-LGBTQI hate speech online. The organisation's activities related to hate speech include implementing different projects, providing legal aid, and raising awareness (e.g. booklets and training for professionals). The organisation also participates in the

European Commission's Code of Conduct on Countering Illegal Hate Speech Online monitoring exercises, which cover hate speech not only against the LGBTQI community but on all grounds. This monitoring focuses on contents deemed illegal under national laws and consists of intensive annual monitoring rounds with a simple manual process (searching – reporting – administrating – following up). The content removal rates resulting from the exercise are publicly available.

### *Monitoring online anti-Roma hate speech in Hungary*

The Hungarian, Roma-led community organisation is focused on Roma youth and education. From September 2020 to July 2021, the organisation participated in an international project covering 10 CEE countries, to gain in-depth nation-specific knowledge regarding online hate speech and raise awareness on the issue amongst youth. As a result of the project, 208 cases of anti-Roma online hate speech were collected, and more than 20 peer-to-peer workshops were organised. Based on the results, anti-Roma hate speech in Hungary is mainly triggered by crime-related content. The results also show that racist hate speech is widespread and is regular on the sites and profiles of far-right political actors. The project's results also showed that coded hate speech is prevalent online.

### **Group discussion**

The discussion following the presentations mainly focused on the methodologies used for monitoring hate speech online and the challenges of monitoring.

### *Methodology*

- As the group required, the four speakers shared the methodology their organisation has used for monitoring in detail.
- Phiren Amenca briefly introduced the monitoring method used in the CHAD project. The project uses artificial intelligence for monitoring social media.

### *Challenges*

- One of the challenges that more participants mentioned is the context in which the hate speech appears, which defines, in many cases, whether a particular content is or is not hate speech.
- Another challenge mentioned was the use of coded language: often, hate speech appears as a coded language, making it more challenging to monitor.